

(D) The responsibilities of institutions conducting publicly funded research

We must develop regulations governing the management of articles, research data, and other research results by organizations equipped with rules, personnel, and equipment for conducting publicly funded research. In particular, we must institute specific policies for preventing the loss, dismantling, or destruction of the results of research activities.

More specifically, all research results and resources should be assigned a Persistent Object Identifier, and a framework for managing these identifiers must be established.

(Reference) Advantages and disadvantages of data sharing

Advantages of data sharing	Disadvantages of data sharing
✓ Fewer errors and mistakes	✓ Concerns regarding improper use
✓ Increased return on investment for research funding	✓ Limited time and personnel resources
✓ Respect for rules of journals and supporting organizations	✓ Costs associated with data expiration dates and data sharing
✓ Less duplication and bias	✓ Concerns regarding the privacy of medical data and related data
✓ Greater reproducibility and validation of research	✓ Uncertainties regarding data ownership and data disclosure restrictions
✓ Further testing of hypotheses	✓ Insufficient incentives and recognition in the academic world
✓ Usage for educational purposes	✓ Insufficient repositories or insufficient awareness of repositories
✓ Merging with other data sets	✓ Protection of confidential corporate data
✓ More citations	

3. On policies for implementing open science initiatives at key institutions

(A) General observations

To ensure the greatest possible access to all users, government agencies, institutions responsible for the distribution of research funds, and universities and other research institutions must cooperate as necessary to develop plans for increasing the effectiveness of these efforts. The following elements may be regarded as essential components of open science implementation plans and policies to be established by such institutions:

- Initiatives to strengthen competitiveness and innovativeness
- Processes that are transparent to stakeholders
- Steps to ensure that digital research data arising from publicly funded research is accessible and located in specific locations
- Steps to guarantee long-term oversight of research results, and methods for optimizing search, storage, and dissemination capabilities to ensure that data is easily accessible and interoperable
- Establishment of guidelines to clarify the obligations of researchers receiving public research funds, including any necessary revisions to associated regulations
- Assessments of the extent to which institutions are honoring their commitments to plans and other agreements, and, as necessary, strategies to ensure the thorough execution of these plans

- Securement of operational expenses and other costs to be incurred in executing plans and other agreements
- Establishment of a road map detailing the priority with which plans are to be executed
- Plans for building data infrastructures (repositories and similar facilities)

(B) Policies regarding access to journals

With regard to the transition to open access to journals, we support open access following the April 2002 manifesto of the Budapest Open Access Initiative. Thus, all articles, including publications and other documents, arising from publicly funded research must be available for all users to access, search, read, and analyze. In addition, long-term maintenance for articles must be provided.

- The planning statements and other documentation prepared by participating institutions must clearly include the following points:
 - Steps for preventing improper large-scale distribution of academic publications, within the feasible range of execution, shall be demonstrated.
 - Steps will be taken to improve the ease of releasing, searching, downloading, analyzing, and accessing material in digital formats.
 - As a general guideline regarding the possibility of using articles, in cases where immediate open access is difficult, it is desirable to establish an appropriate embargo period. In this case, each institution may conduct independent initiatives with due consideration paid to specific research areas, specific research problems, and national interest. In addition, in specific fields of research, a framework shall be prepared for users to request modification of embargoes in cases where is not possible to establish alignment with plans.
 - After the end of an embargo, the article will be provided in its entirety. It is also desirable to provide addition to the data that illustrate the main points of the article, as well as access to any relevant supporting material.
 - The rights assigned to authors, journals, and publishers shall be preserved.
- Publications and metadata will be securely preserved with the following conditions satisfied:
 - Long-term maintenance of content, and cost-free access to that content, shall be provided.
 - Text and auxiliary content (graphics, images, supplementary data) will be supplied in standard, widely-used formats.
- Repositories must demonstrate support for Green open access (Green OA) and Gold open access (Gold OA).
- In particular, with regard to a mandatory transition to open access premised on the transition to Green OA for published results, policies should be enacted with due consideration paid to the fact that a global consensus is emerging among global participants.

(C) Access to digitized research data

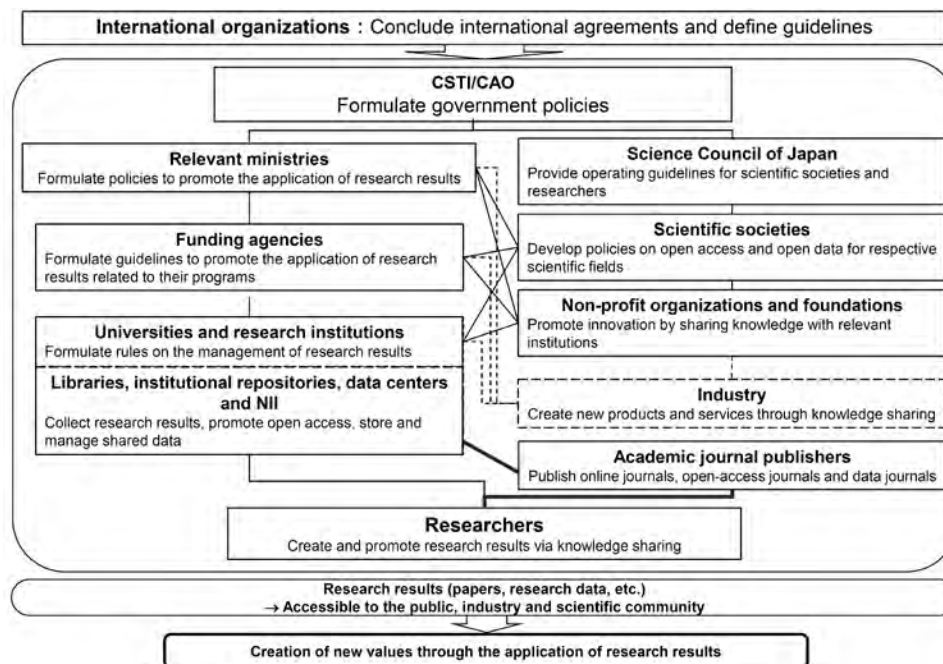
Plans and other documents must be formulated in a way that accounts for the particular characteristics of various research fields, and which recognizes that differences exist in methods used for storing and sharing research data. These plans must ensure that the results of publicly funded research, together with any digitally formatted data arising from this research, are preserved in such a way as to allow use, including access, search, retrieval, and analysis, by any user when made publicly available.

- It is essential to maximize cost-free access to research data stored in digital formats to all users. However,
 - Confidentiality and individual privacy must be protected.
 - Results or by-products for which property rights exist must be protected.
 - A road map will be prepared that lays out a data infrastructure allowing long-term storage.
- This road map must be accompanied by usage rules demonstrating that the data provided may be freely used.
 - Because research data are not authored content, it is desirable to adopt a CC0 (Creative Commons Zero). Moreover, with regard to matters such as quality control and the aggregation and organization of data, in cases involving copyrights in the form of intellectual property, such as databases reflecting the labor and sophisticated expertise of expert researchers, it is desirable to adopt a CC-BY (Creative Commons Attribution, BY).
- Researchers receiving support from public research funds will prepare data management plans as necessary. In cases in which it is difficult to ensure long-term storage of or access to data, these plans will also be used to explain the reasons for the difficulties.
- Data management plans submitted by researchers will be properly assessed.
- A framework shall be established for verifying that researchers are adhering to their own data management plans.
- In appropriate cases, users will be encouraged to entrust data to accessible databases.
- With regard to research datasets that may be used within a given plan, appropriate attributes will be specified, and a platform for providing these datasets will be developed.
- Personnel development, training, and education in the use of technologies for managing, analyzing, maintaining, storing, and sharing research data shall be supported.

(D) Execution and follow-up for plans to implement open science initiatives

- When an institution has formulated a plan or similar document, efforts must be made to ensure that this document is widely communicated, for example, by posting the document on the institution's website.
- The Cabinet Office, Government of Japan, as well as Japan's Council for Science, Technology, and Innovation, shall periodically check the status of execution of each institution's plans and other initiatives.
 - The establishment and state of progress of data management infrastructure will be monitored.
 - With an eye toward stimulating the use of data for various applications, follow-up monitoring will be conducted to assess the status of matters such as platform architecture, storage and sharing technologies, and personnel development.

Correlation diagram of policy making and implementation



4. Issues to consider in the promotion of open science

In advancing the cause of open science, it is essential to pay careful attention to a number of issues that have been raised. For this reason, we expect each institution to consider the following points when establishing their policies and plans for implementing open science initiatives.

(A) Differing definitions of “free or gratis” and “libre”

With regard to open access, it has been noted that some confusion exists between the definitions of the terms “cost-free” and “freely available” among participants. It is important to clarify the definition of these terms in an effort to promote a common understanding among all participants.

(B) Copyright policies

It is essential to address certain issues related to copyrights which cannot be avoided in promoting the use of research results. For example, according to the Society Copyright Policies in Japan (SCPJ), among 2,609 registered academic societies, 1,355 do not have copyright policies (or have non-publicly-disclosed policies, or did not respond) (as of March 20, 2015).

(C) The definition of ownership for large-scale datasets

It is essential to address certain problems in relation to the definition of ownership for large-scale datasets that may be compiled by providers of computers or software, particularly regarding privacy, confidentiality, and security for medical records, vital records, and other specific classes of data.

(D) Establishing a relationship of coexistence with the academic publishing world

It is essential to determine an organized response to entities in the academic publishing world engaged in developing new business models and providing new services of data curation and storage for open access. This includes rules of possession for public intellectual property: are the data entrusted to a publishing company or another third-party organization, or are they controlled internally (in Japan)?

(E) Incentives for researchers and the scientific community

For individual researchers, offering worldwide access to research data can not only enhance the value of individual results, deepening or extending findings and thus contributing to the advancement of science itself, but also heighten the societal meaningfulness of research endeavors. However, it is essential to establish protocols to avoid policies that allow specific researchers to gain priority access to all information. The culture of secrecy in the research world, and the competitive environment surrounding unilateral public releases, can impede collaboration among researchers; this may serve to discourage researchers from sharing or expanding access to their ideas and research data. In the past, activities such as creating data and making it available for other researchers have gone underappreciated in many fields relative to the publishing of research papers. For this reason, it is important to use policy inducements to strengthen incentives for researchers and the scientific community, ensuring that efforts to promote openness are properly recognized.

(F) Accommodating the characteristics of particular research fields

It is essential to recognize the differences that exist among research fields (e.g., physics, chemistry, material science, earth science, biosciences, humanities and social sciences, and more) in methods used for storing and sharing research data, and to construct rules appropriate for the particular characteristics of each field. We must appreciate the wide variety of specialized circumstances that may be present in particular fields. These include the following items: (a) the number of participants who understand data sharing and openness within a given field of research; (b) the technical ease or difficulty of handling individual data sets; and (c) the relationship between the need for specialized skills and research abilities and the magnitude of the efforts required of researchers to generate and organize data. We must also recognize the fact that, even when working with sophisticated research data, in some research fields, the following two types of data coexist: (a) data gathered from novel experiments and equipment by universities and research institutions for research and development purposes in fields such as satellite-based observatories or seismic and meteorological science; and (b) data that may be organized and publicly released for administrative purposes.

(G) The skills needed for open science

Not all researchers possess the skills needed to share their articles and data online in an open fashion. Although it is common practice in some fields of research to accumulate large quantities of research data, most researchers will require training and education to develop the skills needed to realize the imperatives of open science.

(H) Technical infrastructure and personnel development

Metadata that describe research data, algorithms that process data, and the organization of other elements of online infrastructure are essential tools for open science. Interoperability between

different infrastructures can be a barrier to the sharing of data. Moreover, to ensure effective reuse of data, it is essential to provide quality control for data and to associate data with appropriate metadata. In universities and other research institutions, it is essential to investigate and promote comprehensive training systems to enable the establishment of data control systems with key roles played by technical employees, University Research Administrators (URAs), and university library staff. In addition, data scientists and data curators may be retained as resources to provide research support.

(I) Appropriate and sustainable models for the provision of funds

To ensure a robust implementation of open access, researchers at universities and other research institutions must fulfill their obligations regarding open access, and to this end, it is essential to secure funds for management and monitoring operations. Systems must be designed to determine whether these funds should be included in public research funds or prepared separately as part of a scheme to supply supportive funding for the maintenance of open access.

5. Open questions for future study

Efforts to promote open science must grapple with a large number of issues that must be addressed from a long-term perspective. In particular, the following points require ongoing study and consideration.

We expect the Cabinet Office and the Council for Science, Technology, and Innovation to conduct strategic and sustained reviews appropriate for the status of each institution's engagement in initiatives.

(A) Considerations related to the open release and sharing of articles and research data

- It is essential to take steps to ensure that the progress of science and technology is not impeded.
- It is also important to take the perspective of stimulating innovations in private-sector corporations.
- The question of what data is to be shared, including strategic intellectual property, must be comprehensively investigated.
- Initiatives to encourage open disclosure of copyright policies in academic societies
- Systematization of the funding needed for operation and promotion

(B) Considerations related to the storage of research data

Research data must be stored and organized even after the research is completed.

- Frameworks (including data repositories) for the storage and organization of data are required.
- To establish responsibility for shouldering the burden of permanent, long-term storage of articles and research data, it will be useful to enlist the cooperation of libraries, information platform centers, and other institutions that work with scientific and technical resources, academic documents, and archives. Particular examples of such institutions include Japan's National Diet Library, university libraries, and information platform centers.
- Usage and access rights (licensing) for publicly released data and information must be made clear. In addition, there must be discussions of constructing systems for, and establishing a framework to oversee, services for permanent storage and provision of resources.

- We must keep in mind that storage-related expenses, including not only computer-related expenses, but also personnel costs associated with maintaining and preserving data, are important. Steps for reducing costs, such as the use of cloud computing as a computational resource, must be considered.
- To ensure the establishment of open science systems that will remain sustainable into the future, we should invent new business models such as licensing the organized use of data; the image to keep in mind here is the copyright system.

(C) Consideration of which data to store and for what duration

- We must keep in mind that it is unrealistic to store all data.
- It is important for Japan to maintain a presence within the global community for determining rules.
- We must also consider our philosophy regarding “data publishing.”

(D) Considerations regarding assessments of the technical quality of research data and related matters

- We must establish consensus regarding assessments of the technical quality of research data and assessments in accordance with the unique standards of individual fields of research. A framework similar to the peer-review system for such assessments is necessary.

(E) Considerations regarding incentives for researchers and related matters

- Incentives for researchers are essential. For example, researchers who provide high-quality data may be rewarded with appropriate compensation, such as funding assistance or opportunities for promotion. We must consider the establishment of a framework for providing rewards commensurate with results.

(F) Considerations regarding retention of personnel for planning, development, and operation for services to support data-driven research

- The explosive growth in the volume of research data, and the concomitant changes in storage technologies, that have accompanied the modern evolution of ICTs
- Training and utilization of data scientists and similar professionals
- Training and retention of personnel, including data scientists and data curators, capable of understanding the attributes, management techniques, users (with or without specialized skills), and applied aspects of data, including the various differences that exist between fields of science and technology.

Conclusions

In this report, we have considered *open science*, a movement which is rapidly becoming mainstream around the world. We have discussed the ways in which Japanese stakeholders can extract maximum benefit from the advantages of open science while describing basic policies that can effectively mitigate or eliminate various anxieties and issues associated with the transition to a more open world.

In the future, we look forward to the formation of an executable framework based on these basic policies for a robust implementation of the open science paradigm. This will include the establishment of policies and plans for executing open science initiatives to enhance Japan's performance in the fields of scientific and technological innovation, to encourage active collaboration with foreign countries, to maintain Japan's international competitiveness, and to develop a shared recognition of the value of open science among the various Japanese stakeholders.

For this reason, our recommendation to the Cabinet Office and to Japan's Council for Science, Technology, and Innovation is to recognize that the transition to open science is a process that will not be completed overnight. We hope that the goals of open science will be clearly reflected in the Fifth Science and Technology Basic Plan, and we look forward to a robust and ongoing debate regarding any long-term questions that remain to be investigated, while encouraging stakeholders to advance the causes of open science and conducting periodic follow-up studies regarding the status of execution plans developed by various institutions.

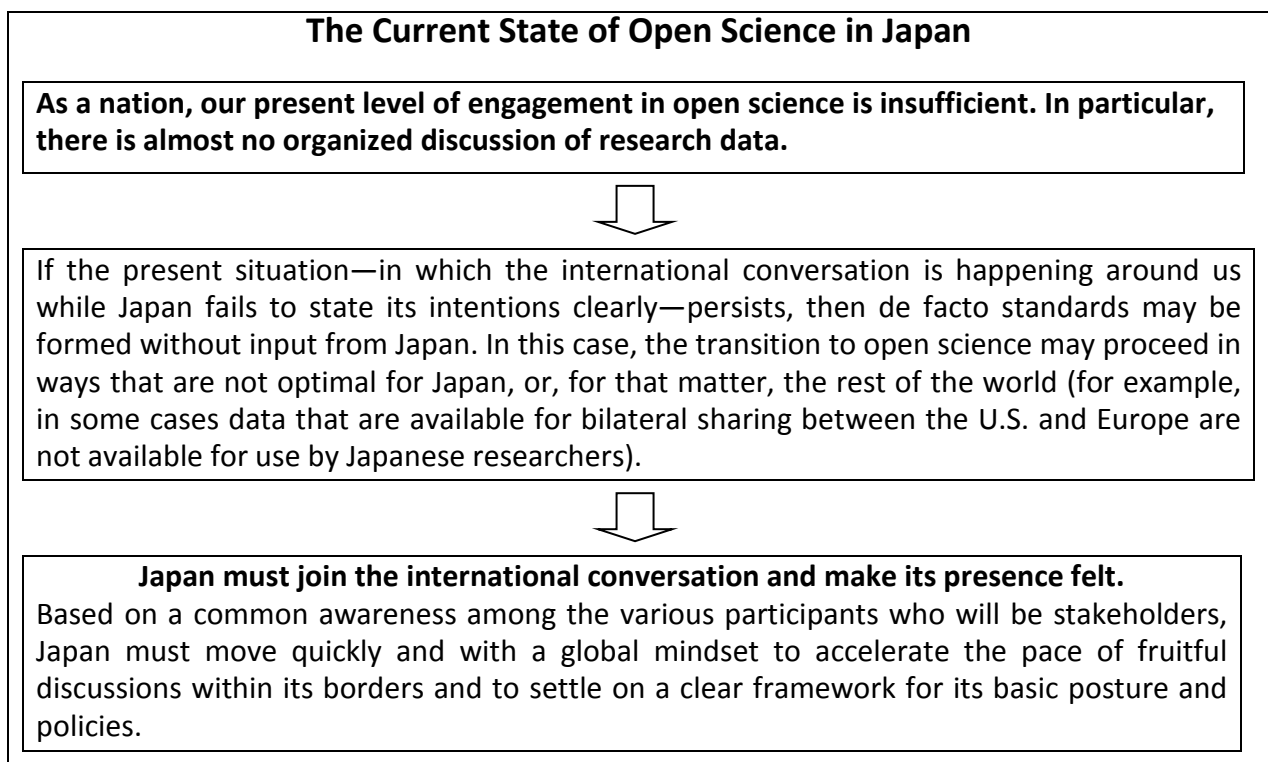
In closing, we hope that the occasion of this report will help to ensure Japan's presence at the dawn of a revolutionary new era in science, one in which new scientific methods will be disseminated and established, in which disparate research fields will merge and international collaborations will flourish, and in which a new science capable of producing value well beyond the range of what was possible in the past will emerge. We have every hope that promoting open science in Japan will help to ensure a bright future for our country.

Appendices

Appendix 1. The current state of open science in Japan

As we noted in the Introduction, the international movement toward increasingly open science shows no signs of abating; today, open access to research data has joined open access to published results as a primary focus of discussion and promotion. For these reasons, it is essential that Japan (a) acquire a thorough understanding of the status of past open science initiatives in Japan, (b) recognize the fields of research in which open sharing of data on a global scale has already begun to progress in earnest, and (c) join the international discussion currently taking place on a worldwide scale to ensure that Japan's presence is felt.

Indeed, by falling behind global trends in the field of open science, Japan risks falling victim to many unfortunate consequences which have been widely pointed out and which we must take pains to keep in mind.



(A) The status of initiatives in Japan

(i) Open access (to articles and journals) and institutional repositories

In Japan, the following two documents called for the advancement of open access through the establishment of institutional repositories and the development of open access journals: (a) the Fourth Science and Technology Basic Plan enacted by Japan's Cabinet in August 2011; and (b) a summary report prepared in July 2012 by the Science Information Infrastructure Working Group of the Council for Science and Technology within Japan's Ministry of Education, Culture, Sports, Science, and Technology (MEXT).

Promotion Policies in the 4th Science and Technology Basic Plan

4. Formation of research environments and platforms comparable to international standards

(3) Establishment of research information platforms

Promotion policies

- The national government will promote the establishment of institutional repositories (see below) at universities and public research institutions, and will encourage educational research institutions to offer systematic collection, storage, and open access for articles, measurements, experimental data, and other resources. In addition, the national government will promote the digitization of academic journals published by academic societies, and will encourage digitization and open access for literature and resources, including work in the humanities and social sciences, possessed by organizations such as the National Diet Library and university libraries.

Institutional repository: An electronic archive system maintained by individual institutions to provide storage and public access for articles and other resources.

- The national government will promote the automation of unified search, structuralization, and knowledge extraction in a way that cuts across boundaries between fields; this includes the networking of digital information resources, data standardization, a basic information infrastructure to indicate the ownership of content, and improved procedures for the association of information. In addition, a *knowledge infrastructure* system will be designed and rolled out to facilitate the search and identification of all research information in a unified fashion.
- The national government expects universities and public research institutions to investigate effective policies for ensuring efficient and stable subscriptions to electronic journals. The national government will assist in supporting such initiatives.

Concrete policies to promote this agenda within MEXT included the following: (a) efforts to promote systematic improvements in Grant-in-Aid for Publication of Scientific Research Results in Grant-in-Aid for Scientific Research; and (b) the creation, in fiscal year 2013, of a new category to support the development of open access journals.

The Japan Science and Technology Agency (JST) has sought to improve the distribution of open access journals in Japan by improving functionality for J-STAGE (a platform to support the distribution of electronic journals). In addition, in April 2013, JST announced that it would encourage open access for the results of JST-assisted research, and that it would explore the possibility of making this mandatory.

The status of open access in Japan

Specific promotion policies

- Public access to academic papers published in open access journals (Gold OA)
- Public access to authors' final drafts of academic papers via repositories (Green OA)

Current Trends in Japan

JSPS: Amended system for Grant-in-Aid for Scientific Research (Grant-in-Aid for Publication of Scientific Research Results)

JST: Assistance via an electronic journal platform (J-STAGE); promoting the assignment of digital object identifiers (DOIs) to guarantee permanent access to academic knowledge; currently considering updating the status of open access to the results of supported research from "recommended" to "mandatory"

NII: Assistance in designing institutional repositories (including JAIRO Cloud)

MEXT: Amending PhD graduation requirements to require PhD theses to be available via the Internet (typically in institutional repositories), thus mandating public access

(Prepared based on resources from MEXT)

Amendments to the system for providing Grant-in-Aid for Scientific Research (Grant-in-Aid for Publication of Scientific Research Results)

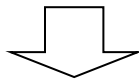
(The areas in which the system needs improvement)

- Support for the expenses needed to publish journals
- Evaluations regarding the content of initiatives to strengthen Japan's ability to communicate information to the rest of the world
- Support for open access initiatives

Before 2012:

Grant-in-Aid for Scientific Research (Grant-in-Aid for Publication of Scientific Research Results): Published academic periodicals

Support for the publishing of journals periodically produced by academic societies on *paper media*



Since 2013:

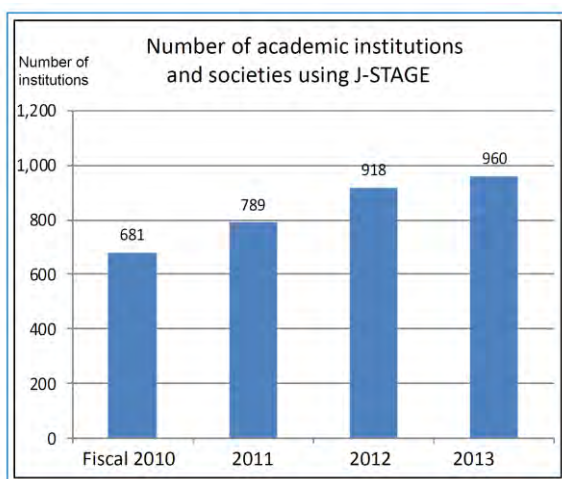
Grant-in-Aid for Scientific Research (Grant-in-Aid for Publication of Scientific Research Results): Strengthening the international communication of information

Support for expenses incurred in *initiatives to strengthen Japan's ability to communicate information to the rest of the world* (peer-review surveys, editing, publishing, and distribution via electronic journals)

(Prepared based on resources from MEXT)

J-STAGE (Japan Science and Technology Information Aggregator, Electronic)

- An electronic journal platform provided to academic institutions and societies by the Japan Science and Technology Agency (JST)
- Currently includes some 1,700 journals from some 960 academic societies in Japan; of these, 87% offer free access
- Registration to J-STAGE is free (excepting certain premium options)



(Prepared based on resources from MEXT)

The National Institute of Informatics (NII) has promoted the establishment of institutional repositories at universities and similar institutions, and has made various efforts to stimulate the distribution of academic content, including a common platform named JAIRO Cloud. NII also convenes seminars that seek to increase understanding of open access, and hosts the International Scholarly Communication Initiative (SPARC Japan). NII amended its graduation requirements in fiscal year 2013 to mandate that PhD theses be made available via the Internet. As the establishment of institutional repositories at universities and similar institutions proceeds, the increased quality of the articles, and the greater opportunities to make use of them, has increased awareness of the key role of open access. Still, overall, initiatives to ensure open access cannot yet be said to have taken root to an acceptable degree.

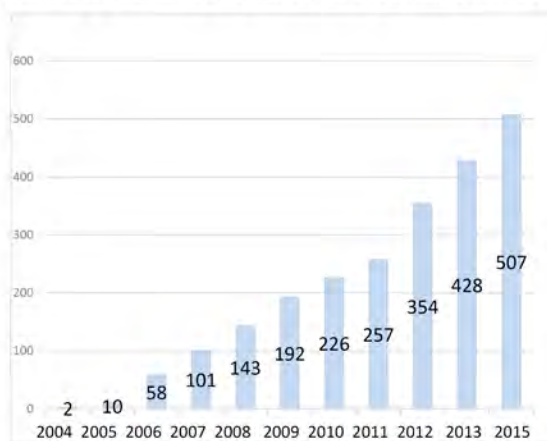
Regarding repositories for research data, at present, the Institutional Repository Promotion Committee of NII is considering research data management policies. Their activities include evaluations of existing data repository systems, experiments involving the storage of data in repositories (policies, systems, and personnel development), and surveys of overseas trends.

Institutional Repositories

Project of the National Institute for Informatics (NII) to support the establishment of institutional repositories

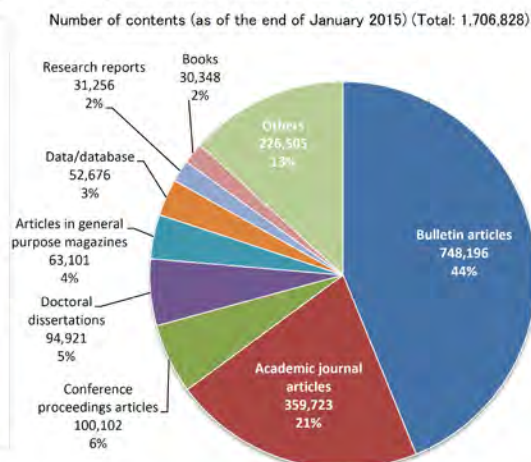
- Infrastructure creation will proceed primarily at national universities; repositories will be used as storage media for communicating the results of education and research in universities and other institutions
- Over one million individual items (educational and research results) will be stored and communicated.

○ Number of institutions that have established public repositories



Source: National Institute for Informatics, Institutional Repositories Program. "Statistics on Institutional Repositories" and "IRDB Content Analysis System"

○ Contents of institutional repositories



Others: Also includes technical reports and resource materials for conference presentations

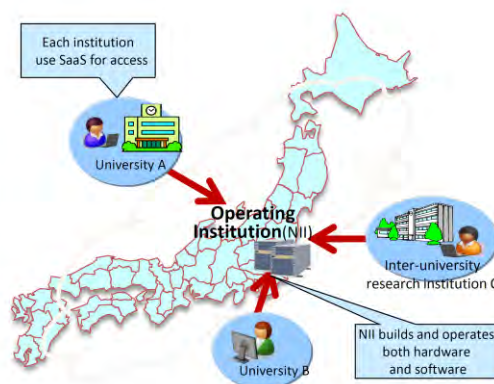
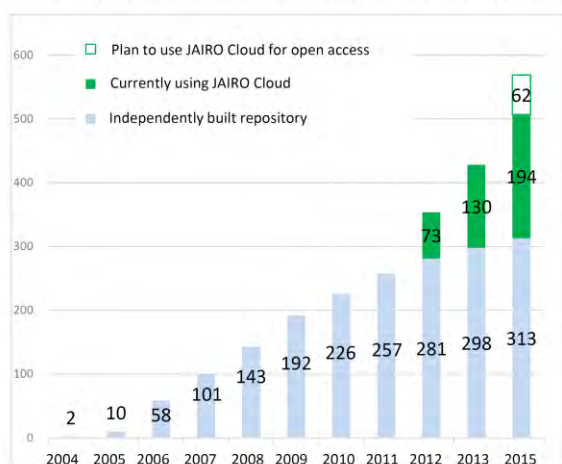
(Prepared based on resources from MEXT)

JAIRO Cloud (Shared repository service)

A cloud service developed and provided by the National Institute (NII) for Informatics to support institutional repositories

In an effort to encourage open access and promote the establishment of institutional repositories at universities and other educational research institutions, this service provides shared repositories for universities that would find it difficult to build or operate their own repositories.

○ The change in the number of institutions offering institutional repositories



A framework that makes it easy for individual institutions to register and offer open access to their content

(Prepared based on resources from MEXT)

(ii) Data sharing and the transition to databases

According to a report from the International Science Data Analysis group of the Information Science Committee of the Science Council of Japan, the U.S., China, and India have between a few dozen and a few hundred database centers in each of various research fields, allowing the development of a large number of databases. Although this type of large-scale database center does not exist in Japan, databases have arisen in Japan within various fields, including research, education, industry, government, and medicine. However, because the nation as a whole does not have a data strategy, there is little opportunity for these databases to achieve organic interconnection. Conversely, they are scattered somewhat randomly, and it has been observed that nobody can understand overall data activities in Japan. In short, Japan's data-utilization infrastructure is shabby compared to that of foreign countries; the primary cause of this is not simply a lack of national competency, but instead the failure of the national government to establish a clearly defined data strategy.

In the field of research in particular, the transition to database structures has proceeded apace in some fields, including the life sciences, earth observation science, and material science, but the question of how to maximize data utilization remains open.

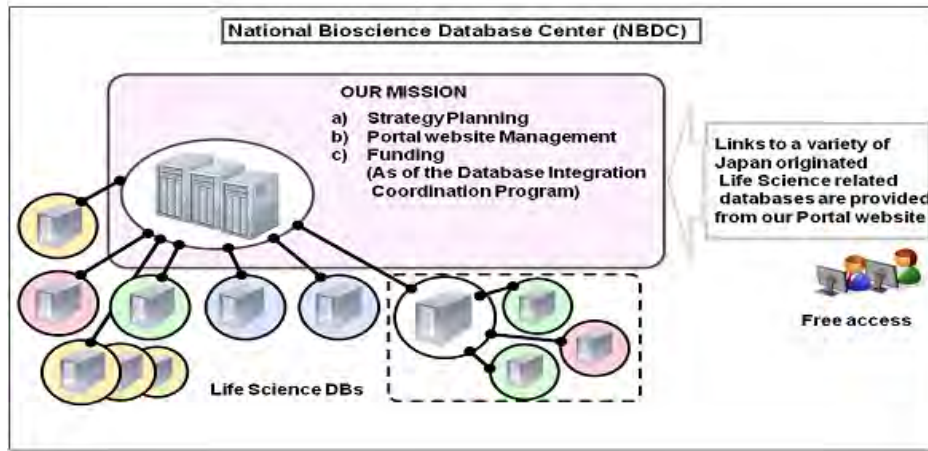
(a) Life sciences

In the field of life sciences, the Council for Science and Technology Policy (today the Council for Science, Technology, and Innovation), following policies indicated by the (previously known as) Comprehensive Data Task Force of its Life Sciences Project Team, has promoted the objective of database consolidation for the life sciences at the JST's National Bioscience Database Center (NBDC), which was established in April 2011.

By consolidating databases in the life sciences, and thus maximizing the value of data, the NBDC aspires to become a database center that contributes to the productivity of not only Japanese users, but also users worldwide, and of which Japan can boast proudly. The center is actively engaged in promoting the following four activities: (a) Strategic Planning; (b) the Creation and Operation of Portal Sites; (c) the Research and Development of Platform Technologies for Database Consolidation; and (d) the Consolidation of Bio-Related Databases.

However, several problems remain. Among these are the following: (a) the problem of how to design incentives to spur data provision, given that the sense of ownership among data producers remains strong; (b) disagreements regarding semantics and terminology between researchers and fields; and (c) the fact that newly available equipment and facilities are still inadequate given the explosive growth and the increasing diversity of data in the field of life sciences.

Incidentally, in addition to the NBDC, Japan also has the DNA Data Bank of Japan, a database operated by the National Institute of Genetics that stores three-dimensional structures of proteins and gene sequences, and the Protein Data Bank of Japan, an international data bank operated by the Institute for Protein Research at Osaka University. However, adding up the numbers of personnel at all of these centers still yields a number of staff members that is extremely small compared to what one finds in the U.S. and Europe.



(b) Earth science and space science

In the fields of earth science and space science, the World Data Center (WDC) project of the International Council for Science (ICSU) was launched in 1957 and has been active for over 50 years. In addition to operating authorized centers for WDC resources (of which there are 7 in Japan), this organization boasts an extensive track record of activities in data storage and other services directed toward scientists. In recent years, advances in science and technology have spurred the execution of many projects, including new methods of gathering data, new ways to organize databases, and new ways to make data publicly available. However, because there was no standardization of formats, public disclosure rules, grace periods, or rights for data providers, some fields remained in which common international rules could not be applied. Moreover, even though the organization was responsible for obtaining and accumulating high-value data, little attention was paid to considerations such as public availability, data formats and the creation of metadata catalog information to facilitate general data sharing, and the allotment of human resources to organize the data in ways that could be useful to users. For this reason, the organization, in some cases, failed to achieve the maximal opportunity for the use of data.

